USHARAMA
COLLEGE OF ENGINEERING AND TECHNOLOGY
On NH16, Telaprolu, Near Gannavaram, Krishna Dist. – 521109
Ph.: 0866-2527558, www.usharama.edu.in
(An Autonomous Institute, Approved by AICTE, New Delhi, Permanently Affiliated to JNTUK, Kakinada)

Subject Name & Code : Machine Learning (PCIT601)

Exam Name : Q1

1. What is Machine learning? [A]

A) The autonomous acquisition of knowledge through the use of computer programs
B) The autonomous acquisition of knowledge through the use of manual programs
C) The selective acquisition of knowledge through the use of computer programs
D) The selective acquisition of knowledge through the use of manual programs

2. Which of the factors affect the performance of learner system does not include? [D]

A) Representation scheme used
B) Training scenario
C) Type of feedback
D) Good data structures

3. Different learning methods does not include? [D]

A) Memorization
B) Analogy
C) Deduction
D) Introduction

4. Some telecommunication company wants to segment their customers into distinct groups ,this is an example of [C]

A) supervised learning
B) reinforcement learning
C) unsupervised learning
D) data extraction

5. In the example of predicting number of babies based on stork's population ,Number of babies is [A]

A) outcome
B) feature
C) observation
D) attribute

6. What is true about Machine Learning? [D]

A) Machine Learning (ML) is that field of computer science
B) ML is a type of artificial intelligence that extract patterns out of raw data by using an algorithm or method.
C) he main focus of ML is to allow computer systems learn from experience without being explicitly programmed or human intervention
D) All of the mentioned

7. ML is a field of AI consisting of learning algorithms that? [D]

A) Improve their performance
B) At executing some task
C) Over time with experience
D) All of the mentioned

8. The action _____ of a robot arm specify to Place block A on block B. [A]

A) STACK(A,B) B) LIST(A,B) C) QUEUE(A,B) D) ARRAY(A,B)

9. A_____ begins by hypothesizing a sentence (the symbol S) and successively predicting lower level constituents until individual preterminal symbols are written. [C]

A) bottow-up parser B) top parser C) top-down parser D) bottom parser

10. A model of language consists of the categories which does not include _____. [B]

A) System Unit B) structural units. C) data units D) empirical units

11. The model will be trained with data in one single batch is known as ? [C]

A) Batch learning B) Offline learning C) Batch & Offline learning D) None of the mentioned

12. Which of the following are ML methods? [A]

A) based on human supervision B) supervised Learning C) semi-reinforcement Learning D) All of the mentioned

13. In Model based learning methods, an iterative process takes place on the ML models that are built based on various model parameters, called ? [C]

A) mini-batches B) optimizedparameters C) hyperparameters D) superparameters

14. What characterize is hyperplance in geometrical model of machine learning? [B]

A) a plane with 1 dimensional fewer than number of input attributes B) a plane with 2 dimensional fewer than number of input attributes C) a plane with 1 dimensional more than number of input attributes D) a plane with 2 dimensional more than number of input attributes

15. To find the minimum or the maximum of a function, we set the gradient to zero because: [A]

A) The value of the gradient at extrema of a function is always zero B) Depends on the type of problem C) The value of the gradient at extrema of a function is always zero and Depends on the type of problem D) None of the mentioned

16. Which of the following techniques would perform better for reducing dimensions of a data set? [A]

A) removing columns which have too many missing values B) removing columns which have high variance in data C) removing columns with dissimilar data trends D) None of the mentioned

17. How do you handle missing or corrupted data in a dataset? [D]

A) Drop missing rows or columns B) Replace missing values with mean/median/mode C) Assign a unique category to missing values D) All of the mentioned

18. When performing regression or classification, which of the following is the correct way to preprocess the data? [A]

A) Normalize the data -> PCA -> training B) PCA -> normalize PCA output -> training C) Normalize the data -> PCA -> normalize PCA output -> training D) None of the mentioned

19. Database query is used to uncover this type of knowledge. [D]

A) deep B) hidden C) shallow D) multidimensional

20. Which of the following techniques can not be used for normalization in text mining? [C]

A) Stemming B) Lemmatization C) Stop Word Removal D) None of the mentioned

21. In which of the following cases will K-means clustering fail to give good results? [D]
1) Data points with outliers
2) Data points with different densities
3) Data points with nonconvex shapes

A) 1 and 2 B) 2 and 3 C) 1 and 3 D) All of the mentioned

22. Which of the following is a reasonable way to select the number of principal components "k"? [A]

A) Choose k to be the smallest value so that at least 99% of the varinace is retained. B) Choose k to be 99% of m (k = 0.99*m, rounded to the nearest integer). C) Choose k to be the largest value so that 99% of the variance is retained. D) Use the elbow method.

23. What is a sentence parser typically used for? [B]

A) It is used to parse sentences to check if they are utf-8 compliant. B) It is used to parse sentences to derive their most likely syntax tree structures. C) It is used to parse sentences to assign POS tags to all tokens. D) It is used to check if sentences can be parsed into meaningful tokens.

24. A machine learning problem involves four attributes plus a class. The attributes have 3, 2, 2, and 2 possible values each. The class has 3 possible values. How many maximum possible different examples are there? [D]

A) 12 B) 24 C) 48 D) 72

25. Application of machine learning methods to large databases is called [A]

A) data mining. B) artificial intelligence C) big data computing D) internet of things

26. If machine learning model output involves target variable then that model is called as [B]

A) descriptive model B) predictive model C) reinforcement learning D) All of the mentioned

27. Father of Machine Learning (ML) [C]

A) Geoffrey Chaucer B) Geoffrey Hill C) Geoffrey Everest Hinton D) None of the mentioned

28. Which is FALSE regarding regression? [C]

A) It may be used for interpretation B) It is used for prediction C) It discovers causal relationships D) It relates inputs to outputs

29. Choose the correct option regarding machine learning (ML) and artificial intelligence (AI) [D]

A) ML is a set of techniques that turns a dataset into a software B) AI is a software that can emulate the human mind C) ML is an alternate way of programming intelligent machines D) All of the mentioned

30. In general, to have a well-defined learning problem, we must identity which of the following [D]

A) The class of tasks B) The measure of performance to be improved C) The source of experience D) All of the mentioned

31. Successful applications of ML [D]

A) Learning to recognize spoken words B) Learning to drive an autonomous vehicle C) Learning to classify new astronomical structures D) All of the mentioned

32. Which of the following does not include different learning methods [B]

A) Analogy B) Introduction C) Memorization D) Deduction

33. In language understanding, the levels of knowledge that does not include? [A]

A) Empirical B) Logical C) Phonological D) Syntactic

34. Designing a machine learning approach involves:- [D]

A) Choosing the type of training experience B) Choosing the target function to be learned C) Choosing a representation for the target function D) All of the mentioned

35. Concept learning inferred a valued function from training examples of its input and output. [C]

A) Decimal B) Hexadecimal C) Boolean D) All of the mentioned

36. Which of the following is not a supervised learning? [B]

A) Naive Bayesian B) PCA C) Linear Regression D) Decision Tree Answer

37. [C]

What is Machine Learning?

Artificial Intelligence
Deep Learning
Data Statistics

A) Artificial Intelligence B) Deep Learning C) Deep Learning & Artificial Intelligence D) All of the mentioned

38. What kind of learning algorithm for “Facial identities or facial expressions”? [B]

A) Prediction B) Recognition Patterns C) Generating Patterns D) Recognizing Anomalies Answer

39. Real-Time decisions, Game AI, Learning Tasks, Skill Aquisition, and Robot Navigation are applications of which of the folowing [B]

A) Supervised Learning: Classification B) Reinforcement Learning C) Unsupervised Learning: Clustering D) Unsupervised Learning: Regression

40. Targetted marketing, Recommended Systems, and Customer Segmentation are applications in which of the following [C]

A) Supervised Learning: Classification B) Reinforcement Learning C) Unsupervised Learning: Clustering D) Unsupervised Learning: Regression

41. Fraud Detection, Image Classification, Diagnostic, and Customer Retention are applications in which of the following [A]

A) Supervised Learning: Classification B) Reinforcement Learning C) Unsupervised Learning: Clustering D) Unsupervised Learning: Regression

42. Which of the following is not function of symbolic in the various function representation of Machine Learning? [C]

A) Rules in propotional Logic B) Hidden-Markov Models (HMM) C) Rules in first-order predicate logic D) Decision Trees

43. Which of the following is not numerical functions in the various function representation of Machine Learning? [C]

A) Neural Network B) Support Vector Machines C) Case-based D) Linear Regression

44. FIND-S Algorithm starts from the most specific hypothesis and generalize it by considering only [B]

A) Negative B) Positive C) Negative or Positive D) None of the mentioned

45. FIND-S algorithm ignores [A]

A) Negative B) Positive C) Negative or Positive D) None of the mentioned

46. The Candidate-Elimination Algorithm represents the [B]

A) Solution Space B) Version Space C) Elimination Space D) All of the mentioned

47. Inductive learning takes examples and generalizes rather than starting with [B]

A) Inductive B) Existing C) Deductive D) None of the mentioned

48. Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging? [B]

A) Decision Tree B) Random Forest C) Regression D) Classification

49. o find the minimum or the maximum of a function, we set the gradient to zero because which of the following [B]

A) Depends on the type of problem B) The value of the gradient at extrema of a function is always zero C) Depends on the type of solution D) All of the mentioned

50. Which of the following is a disadvantage of decision trees? [A]

A) Decision trees are prone to be overfit B) Decision trees are robust to outliers C) Factor analysis D) None of the mentioned

51. What is perceptron? [A]

A) A single layer feed-forward neural network with pre-processing B) A neural network that contains feedback C) A double layer auto-associative neural network D) An auto-associative neural network

52. What is Neuro software? [C]

A) It is software used by Neurosurgeon B) Designed to aid experts in real world C) It is powerful and easy neural network D) A software used to analyze neurons

53. Which is true for neural networks? [D]

A) Each node computes it‟s weighted input B) Node could be in excited state or non-excited state C) It has set of nodes and connections D) All of the mentioned

54. What is the objective of backpropagation algorithm? [A]

A) To develop learning algorithm for multilayer feedforward neural network, so that network can be trained to capture the mapping implicitly
B) To develop learning algorithm for multilayer feedforward neural network
C) To develop learning algorithm for single layer feedforward neural network
D) All of the mentioned

55. What is true regarding backpropagation rule? [D]

A) Error in output is propagated backwards only to determine weight updates
B) There is no feedback of signal at nay stage
C) It is also called generalized delta rule
D) All of the mentioned

56. An auto-associative network is [B]

A) A neural network that has only one loop
B) A neural network that contains feedback
C) A single layer feed-forward neural network with pre-processing
D) A neural network that contains no loops

57. A 3-input neuron has weights 1, 4 and 3. The transfer function is linear with the constant of proportionality being equal to 3. The inputs are 4, 8 and 5 respectively. What will be the output? [B]

A) 139
B) 153
C) 162
D) 160

58. What of the following is true regarding backpropagation rule? [B]

A) Hidden layers output is not all important, they are only meant for supporting input and output layers
B) Actual output is determined by computing the outputs of units for each hidden layer
C) It is a feedback neural network
D) None of the mentioned

59. What is back propagation? [B]

A) It is another name given to the curvy function in the perceptron
B) It is the transmission of error back through the network to allow weights to be adjusted so that the network can learn
C) It is another name given to the curvy function in the non perceptron
D) None of the mentioned

60. The general limitations of back propagation rule is/are [D]

A) Scaling
B) Slow convergence
C) Local minima problem
D) All of the mentioned

61. What is the meaning of generalized in statement “backpropagation is a generalized delta rule” ? [C]

A) Because delta is applied to only input and output layers, thus making it more simple and generalized
B) It has no significance
C) Because delta rule can be extended to hidden layer units
D) None of the mentioned

62. Neural Networks are complex functions with many parameter [A]

A) Linear B) Non linear C) Discreate D) Exponential

63. the general tasks that are performed with backpropagation algorithm [D]

A) Pattern mapping B) Prediction C) Function approximation D) All of the mentioned

64. In backpropagation rule, how to stop the learning process? [B]

A) No heuristic criteria exist B) On basis of average gradient value C) There is convergence involved D) None of the mentioned

65. Applications of NN (Neural Network) [D]

A) Risk management B) Data validation C) Sales forecasting D) All of the mentioned

66. The network that involves backward links from output to the input and hidden layers is known as [A]

A) Recurrent neural network B) Self organizing maps C) Perceptrons D) Single layered perceptron

67. Which of the following is/are the decision tree nodes? [D]

A) End Nodes B) Decision Nodes C) Chance Nodes D) All of the mentioned

68. End Nodes are represented by which of the following [B]

A) Solar street light B) Triangles C) Circles D) Squares

69. Decision Nodes are represented by which of the following [D]

A) Solar street light B) Triangles C) Circles D) Squares

70. Chance Nodes are represented by which of the following [C]

A) Solar street light B) Triangles C) Circles D) Squares

71. Advantage of Decision Trees [D]

A) Possible Scenarios can be added B) Use a white box model, if given result is provided by a model C) Worst, best and expected values can be determined for different scenarios D) All of the mentioned

72. terms are required for building a bayes model. [C]

A) 1 B) 2 C) 3 D) 4

73. Which of the following is the consequence between a node and its predecessors while creating bayesian network? [A]

A) Conditionally independent B) Functionally dependent C) Both Conditionally dependant & Dependant D) Dependent

74. Why it is needed to make probabilistic systems feasible in the world? [C]

A) Feasibility B) Reliability C) Crucial robustness D) None of the mentioned

75. Bayes rule can be used for:- [C]

A) Solving queries B) Increasing complexity C) Answering probabilistic query D) Decreasing complexity

76. provides way and means of weighing up the desirability of goals and the likelihood of achieving [A]

A) Utility theory B) Decision theory C) Bayesian networks D) Probability theory

77. Which of the following provided by the Bayesian Network? [C]

A) Complete description of the problem B) Partial description of the domain C) Complete description of the domain D) All of the mentioned

78. Probability provides a way of summarizing the that comes from our laziness and [B]

A) Belief B) Uncertaintity C) Joint probability distributions D) Randomness

79. The entries in the full joint probability distribution can be calculated as [C]

A) Using variables B) Both Using variables & information C) Using information D) All of the mentioned

80. Causal chain (For example, Smoking cause cancer) gives rise to:- [A]

A) Conditionally Independence B) Conditionally Dependence C) Conditionally Dependence & Independence D) None of the mentioned

81. The bayesian network can be used to answer any query by using: [B]

A) Full distribution B) Joint distribution C) Partial distribution D) All of the mentioned

82. Bayesian networks allow compact specification of:- [A]

A) Joint probability distributions B) Belief C) Propositional logic statements D) All of the mentioned

83. The compactness of the bayesian network can be described by [B]

A) Fully structured B) Locally structured C) Partially structured D) All of the mentioned

84. Which of the following is false regarding EM Algorithm? [C]

A) The alignment provides an estimate of the base or amino acid composition of each column in the site

B) The column-by-column composition of the site already available is used to estimate the probability of finding the site at any position in each of the sequences

C) The row-by-column composition of the site already available is used to estimate the probability

D) None of the mentioned

85. Naïve Bayes Algorithm is a learning algorithm. [A]

A) Supervised B) Reinforcement C) Unsupervised D) None of the mentioned

86. EM algorithm includes two repeated steps, here the step 2 is [C]

A) The normalization B) The maximization step C) The minimization step D) None of the mentioned

87. Examples of Naïve Bayes Algorithm is/are [D]

A) Spam filtration B) Sentimental analysis C) Classifying articles D) All of the mentioned

88. Naïve Bayes algorithm is based on and used for solving classification problems [A]

A) Bayes Theorem B) Candidate elimination algorithm C) EM algorithm D) None of the mentioned

89. Types of Naïve Bayes Model: [D]

A) Gaussian B) Multinomial C) Bernoulli D) All of the mentioned

90. Disadvantages of Naïve Bayes Classifier: [A]

A) Naive Bayes assumes that all features are independent or unrelated, so it cannot learn the relationship between

B) It performs well in Multi-class predictions as compared to the other

C) Naïve Bayes is one of the fast and easy ML algorithms to predict a class of

D) It is the most popular choice for text classification problems

91. The benefit of Naïve Bayes:- [D]

A) Naïve Bayes is one of the fast and easy ML algorithms to predict a class of

B) It is the most popular choice for text classification problems

C) It can be used for Binary as well as Multi-class

D) All of the mentioned

92. A person trained to interact with a human expert in order to capture their knowledge. [D]

A) knowledge programmer B) knowledge developer C) knowledge engineer D) knowledge extractor

93. Full form of MDL? [A]

A) Minimum Description Length B) Maximum Description Length C) Minimum Domain Length D) None of the mentioned

94. For the analysis of ML algorithms, we need [C]

A) Computational learning theory B) Statistical learning theory C) Statistical learning theory & Computational learning theory D) None of the mentioned

95. PAC stand for [A]

A) Probably Approximate Correct B) Probably Approx Correct C) Probably Approximate Computation D) Probably Approx Computation

96. hypothesis h with respect to target concept c and distribution D , is the probability that h will misclassify an instance drawn at random according to D. [A]

A) True Error B) Type 1 Error C) Type 2 Error D) None of the mentioned

97. What are the area CLT comprised of? [D]

A) Sample Complexity B) Computational Complexity C) Mistake Bound D) All of the mentioned

98. What area of CLT tells “How many examples we need to find a good hypothesis ?”? [A]

A) Sample Complexity B) Computational Complexity C) Mistake Bound D) None of the mentioned

99. What area of CLT tells “How much computational power we need to find a good hypothesis ?”? [B]

A) Sample Complexity B) Computational Complexity C) Mistake Bound D) None of the mentioned

100. What area of CLT tells “How many mistakes we will make before finding a good hypothesis ?”? [C]

A) Sample Complexity B) Computational Complexity C) Mistake Bound D) None of the mentioned

101. How large is the hypothesis space when we have n Boolean attributes? [A]

A) |H| = 3 n B) |H| = 2 n C) |H| = 1 n D) |H| = 4n

102. The VC dimension of hypothesis space H1 is larger than the VC dimension of hypothesis space H2. Which of the following can be inferred from this? [A]

A) The number of examples required for learning a hypothesis in H1 is larger than the number of examples required for H2
B) The number of examples required for learning a hypothesis in H1 is smaller than the number of examples required for
C) No relation to number of samples required for PAC learning
D) None of the mentioned

103. For a particular learning task, if the requirement of error parameter changes from 0.1 to 0.01. How many more samples will be required for PAC learning? [D]

A) Same
B) 2 times
C) 1000 times
D) 10 times

104. Computational complexity of classes of learning problems depends on which of the following? [D]

A) The size or complexity of the hypothesis space considered by learner
B) The accuracy to which the target concept must be approximated
C) The probability that the learner will output a successful hypothesis
D) All of the mentioned

105. The instance-based learner is a [A]

A) Lazy-learner
B) Eager learner
C) Can"t say
D) None of the mentioned

106. When to consider nearest neighbour algorithms? [D]

A) Instance map to point in kn
B) Not more than 20 attributes per instance
C) Lots of training data
D) All of the mentioned

107. What are the advantages of Nearest neighbour alogo? [D]

A) Training is very fast
B) Can learn complex target functions
C) Don"t lose information
D) All of the mentioned

108. What are the difficulties with k-nearest neighbour algo? [C]

A) Calculate the distance of the test case from all training cases
B) Curse of dimensionality
C) Calculate the distance of the test case from all training cases & Curse of dimensionality
D) None of the mentioned

109. What if the target function is real valued in kNN algo? [A]

A) Calculate the mean of the k nearest neighbours
B) Calculate the SD of the k nearest neighbour
C) Non Calculative
D) None of the mentioned

110. What is/are true about Distance-weighted KNN? [C]

A) The weight of the neighbour is considered
B) The distance of the neighbour is considered
C) The Weight and distance of the neighbour is considered
D) None of the mentioned

111. What is/are advantage(s) of Distance-weighted k-NN over k-NN? [C]

A) Robust to noisy training data
B) Quite effective when a sufficient large set of training data is provided
C) Robust and Quite effective
D) None of the mentioned

112. What is/are advantage(s) of Locally Weighted Regression? [C]

A) Pointwise approximation of complex target function
B) Earlier data has no influence on the new ones
C) Pointwise approximation of complex target function & Earlier data has no influence on the new ones
D) None of the mentioned

113. The quality of the result depends on (LWR) [D]

A) Choice of the function
B) Choice of the kernel function K
C) Choice of the hypothesis space H
D) All of the mentioned

114. How many types of layer in radial basis function neural networks? [A]

A) 3
B) 2
C) 1
D) 4

115. The neurons in the hidden layer contains Gaussian transfer function whose output are to the distance from the centre of the neuron. [B]

A) Directly
B) Inversely
C) equal
D) None of the mentioned

116. PNN/GRNN networks have one neuron for each point in the training file, While RBF network have a variable number of neurons that is usually [A]

A) less than the number of training
B) greater than the number of training points
C) equal to the number of training points
D) None of the mentioned

117. Which network is more accurate when the size of training set between small to medium? [A]

A) PNN/GRNN
B) RBF
C) K-means clustering
D) None of the mentioned

118. What is/are true about RBF network? [D]

A) A kind of supervised learning
B) Design of NN as curve fitting problem
C) Use of multidimensional surface to interpolate the test data
D) All of the mentioned

119. Application of CBR [A]

A) Design B) Planning C) Diagnosis D) All of the mentioned

120. What is/are advantages of CBR? [D]

A) A local approx. is found for each test case B) Knowledge is in a form understandable to human C) Fast to train D) All of the mentioned

121. In k-NN algorithm, given a set of training examples and the value of k < size of training set (n), the algorithm predicts the class of a test example to be the. What is/are advantages of CBR? [B]

A) Least frequent class among the classes of k closest training B) Most frequent class among the classes of k closest training C) Class of the closest D) Most frequent class among the classes of the k farthest training examples

122. Genetic algorithm is a [A]

A) Search technique used in computing to find true or approximate solution to optimization and search problem B) Sorting technique used in computing to find true or approximate solution to optimization and sort problem C) Search and Sort D) None of the mentioned

123. GA techniques are inspired by [A]

A) Evolutionary B) Cytology C) Anatomy D) Ecology

124. When would the genetic algorithm terminate? [C]

A) Maximum number of generations has been produced B) Satisfactory fitness level has been reached for the C) Maximum number of generations has been produced & Satisfactory fitness level has been reached for the D) None of the mentioned

125. The algorithm operates by iteratively updating a pool of hypotheses, called the [A]

A) Population B) Fitness C) Health D) None of the mentioned

126. What is the correct representation of GA? [C]

A) GA(Fitness, Fitness_threshold, p) B) GA(Fitness, Fitness_threshold, p, r) C) GA(Fitness, Fitness_threshold, p, r, m) D) GA(Fitness, Fitness_threshold)

127. Genetic operators includes [C]

A) Crossover B) Mutation C) Crossover & Mutation D) None of the mentioned

128. Produces two new offspring from two parent string by copying selected bits from each parent is called [A]

A) Mutation B) Inheritance C) Crossover D) None of the mentioned

129. Each schema the set of bit strings containing the indicated as [D]

A) 0s, 1s B) only 0s C) only 1s D) 0s, 1s, *s

130. 0*10 represents the set of bit strings that includes exactly (A) 0010, 0110 [A]

A) 0010, 0010 B) 0100, 0110 C) 0100, 0010 D) 0000,0110

131. Correct (h) is the percent of all training examples correctly classified by hypothesis then Fitness function is equal to [A]

A) Fitness (h) = (correct (h)) 2 B) Fitness (h) = (correct (h)) 3 C) Fitness (h) = (correct (h)) D) Fitness (h) = (correct (h)) 4

132. evolution over many generations was directly influenced by the experiences of individual organisms during their lifetime [B]

A) Baldwin B) Lamarckian C) Bayes D) None of the mentioned

133. Search through the hypothesis space cannot be characterized. Why? [A]

A) Hypotheses are created by crossover and mutation operators that allow radical changes between successive generations B) Hypotheses are not created by crossover and mutation C) Hypotheses are created by crossover and not a mutation operators that allow radical changes between successive generations D) None of the mentioned

134. ILP stand for [B]

A) Inductive Logical programming B) Inductive Logic Programming C) Inductive Logical Program D) Inductive Logic Program

135. What is/are the requirement for the Learn-One-Rule method? [C]

A) Input, accepts a set of +ve and -ve training examples. B) Output, delivers a single rule that covers many +ve examples and few -ve. C) Input, accepts a set of +ve and -ve training examples. & Output, delivers a single rule that covers many +ve examples & Output rule has a high accuracy but not necessarily a highand few -ve. D) Output rule has a high accuracy but not necessarily a high

136. is any predicate (or its negation) applied to any set of terms [A]

A) Literal B) Null C) Clause D) None of the mentioned

137. Ground literal is a literal that [C]

A) Contains only variables B) does not contains any functions C) does not contains any variables D) Contains only functions Answer

138. emphasizes learning feedback that evaluates the learner's performance without providing standards of correctness in the form of behavioural [A]

A) Reinforcement learning B) Supervised Learning C) UnSupervised Learning D) None of the mentioned

139. Features of Reinforcement learning [D]

A) Set of problem rather than set of techniques B) RL is training by reward C) RL is learning from trial and error with the D) All of the mentioned

140. Which type of feedback used by RL [B]

A) Purely Instructive feedback B) Purely Evaluative feedback C) Non Evaluative D) None of the mentioned

141. What is/are the problem solving methods for RL? [D]

A) Dynamic programming B) Monte Carlo Methods C) Temporal-difference learning D) All of the mentioned

142. The FIND-S Algorithm [A]

A) Starts with starts from the most specific hypothesis Answer B) It considers negative examples C) It considers both negative and positive D) None of the mentioned

143. The Version space is: [A]

A) The subset of all hypotheses is called the version space with respect to the hypothesis space H and the training examples D, because it contains all plausible versions of the target B) The version space consists of only specific C) The version space does not consists of only specific D) None of the mentioned

144. The Candidate-Elimination Algorithm [D]

A) The key idea in the Candidate-Elimination algorithm is to output a description of the set of all hypotheses consistent with the training B) Candidate-Elimination algorithm computes the description of this set without explicitly enumerating all of its C) This is accomplished by using the more-general-than partial ordering and maintaining a compact representation of the set of consistent D) All of the mentioned

145. The hypothesis h1 is more-general-than hypothesis h2 (h1 > h2) if and only if h1≥h2 is true and h2≥h1 is false. We also say h2 is more-specific-than h1 [A]

A) The statement is true B) The statement is false C) We cannot D) None of the mentioned

146. The List-Then-Eliminate Algorithm [A]

A) The List-Then-Eliminate algorithm initializes the version space to contain all hypotheses in H, then eliminates any hypothesis found inconsistent with any training B) The List-Then-Eliminate algorithm not initializes to the version C) The List-Then-Eliminate algorithm not initializes to the space D) None of the mentioned

147. What will take place as the agent observes its interactions with the world? [A]

A) Learning B) Hearing C) Perceiving D) Speech

148. Which modifies the performance element so that it makes better decision?Performance element [C]

A) Performance element B) Changing element C) Learning element D) None of the mentioned

149. Any hypothesis found to approximate the target function well over a sufficiently large set of training examples will also approximate the target function well over other unobserved example is called: [A]

A) Inductive Learning Hypothesis B) Null Hypothesis C) Actual Hypothesis D) None of the mentioned

150. Feature of ANN in which ANN creates its own organization or representation of information it receives during learning time is [B]

A) Adaptive Learning B) Self Organization C) What-If Analysis D) Supervised Learning

151. How the decision tree reaches its decision? [C]

A) Single test B) Two test C) Sequence of test D) No test

152. Supervised learning and unsupervised clustering both require which is correct according to the statement. [C]

A) output attribute. B) hidden attribute. C) input attribute. D) categorical attribute

153. Tree/Rule based classification algorithms generate which rule to perform the classification. [A]

A) if-then. B) then C) do D) Answer

154. What is Gini Index? [A]

A) It is a type of index structure B) It is a measure of purity C) It is a measure of Non-purity D) None of the mentioned

155. What is not a RNN in machine learning? [A]

A) One output to many inputs B) Many inputs to a single output C) RNNs for nonsequential input D) Many inputs to many outputs

156. Which of the following sentences are correct in reference to Information gain? [D]

A) It is biased towards multi-valued attributes B) ID3 makes use of information gain C) The approach used by ID3 is greedy D) All of the mentioned

157. A Neural Network can answer [B]

A) For Loop questions B) what-if questions C) IF-The-Else Analysis Questions D) None of the mentioned

158. Artificial neural network used for [D]

A) Pattern Recognition B) Classification C) Clustering D) All of the mentioned

159. Which of the following are the advantage/s of Decision Trees? [D]

A) Possible Scenarios can be added B) Use a white box model, If given result is provided by a model C) Worst, best and expected values can be determined for different scenarios D) All of the mentioned

160. What is the mathematical likelihood that something will occur? [C]

A) Classification B) Probability C) Naïve Bayes Classifier D) None of the mentioned